

New Methods for 3D Scene Creation Using 2D Images

Kateřina Dařilkov, Marek Zimanyi, Michal Janořek

Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynska Dolina 842 48
Bratislava, Slovakia
darilkova@sccg.sk, zimanyi@sccg.sk, michal.jancosek@isnslovakia.sk

Abstract

Two of multiple view object reconstruction approaches are inspected in this work. The goal is to create and test several functions for processing of two input images of the scene to uncover geometry of the scene. A new algorithm for Maximally Stable Extremal Regions correspondence detection is presented – True Tentative Correspondences - employing the Sideness constraint. The input to the algorithm is a wide baseline image pair and the output is a set where one element consists of eight tentative correspondences between detected regions, these are the best candidates to compute epipolar geometry between images. There are well-known techniques how to compute a fundamental matrix and reconstruct 3D coordinates. Several of them were tested to find the method that is fast and rather precise. Several ideas for future method combination are mentioned.

Keywords

Two view geometry, epipolar geometry, fundamental matrix, eight point method, algebraic error minimization, Tentative correspondence, Maximally stable extremal regions, Local affine frames, True tentative correspondences.

Introduction

The main goal of the recent computer graphics is to find solutions for navigation and cooperation in virtual environments, especially for the developing of virtual 3D cities. This is what the project Virtual Bratislava (VrBa) is about. That is why the president of Slovak Republic, residing in Bratislava, took up the patronage over this project.

The task of the project implies collection of huge amount (gigabytes) of miscellaneous data and measurements; solving the problem of legal use of data belonging to other owners; development of methods for processing and creating the spatial models, new solutions for navigational and cooperation tools resulting from the specifics of Bratislava as a data source.

A useful cooperation was established with the most significant owners of Bratislava data – the municipal council and the company Eurosense. We initialized the conception of a 3D model. The most important Bratislava objects are the selected cultural monuments – the castle, the coronation church, the mausoleum of Chatam Sofer, etc. The methods developed during the APVT project were successfully applied in a European monuments reconstruction and on-line presentation project in the Culture 2000 framework. The resulting webpage is www.vhce.info (was awarded with the EuroPrix 2004 Top Talent Award Quality Seal in Vienna in autumn 2004).

The original navigation environment for multimedia presentations represents a flexible way of viewing 2D, 2.5D and 3D city parts with relatively small amounts of data, for example a database of hundreds of art photographs, interactive panoramas, small VRML models or video presentations of selected objects.

The cooperative solution led to the creation of a conception of empathic avatars, i.e. virtual guides in the city or in the interior of selected monument by written, spoken and motional cooperation or communication. The project, except the pilot study role, should identify the ways of utilization of the technology of virtual cities (Figure 1.). Therefore basic and applied research on local and international level will follow.

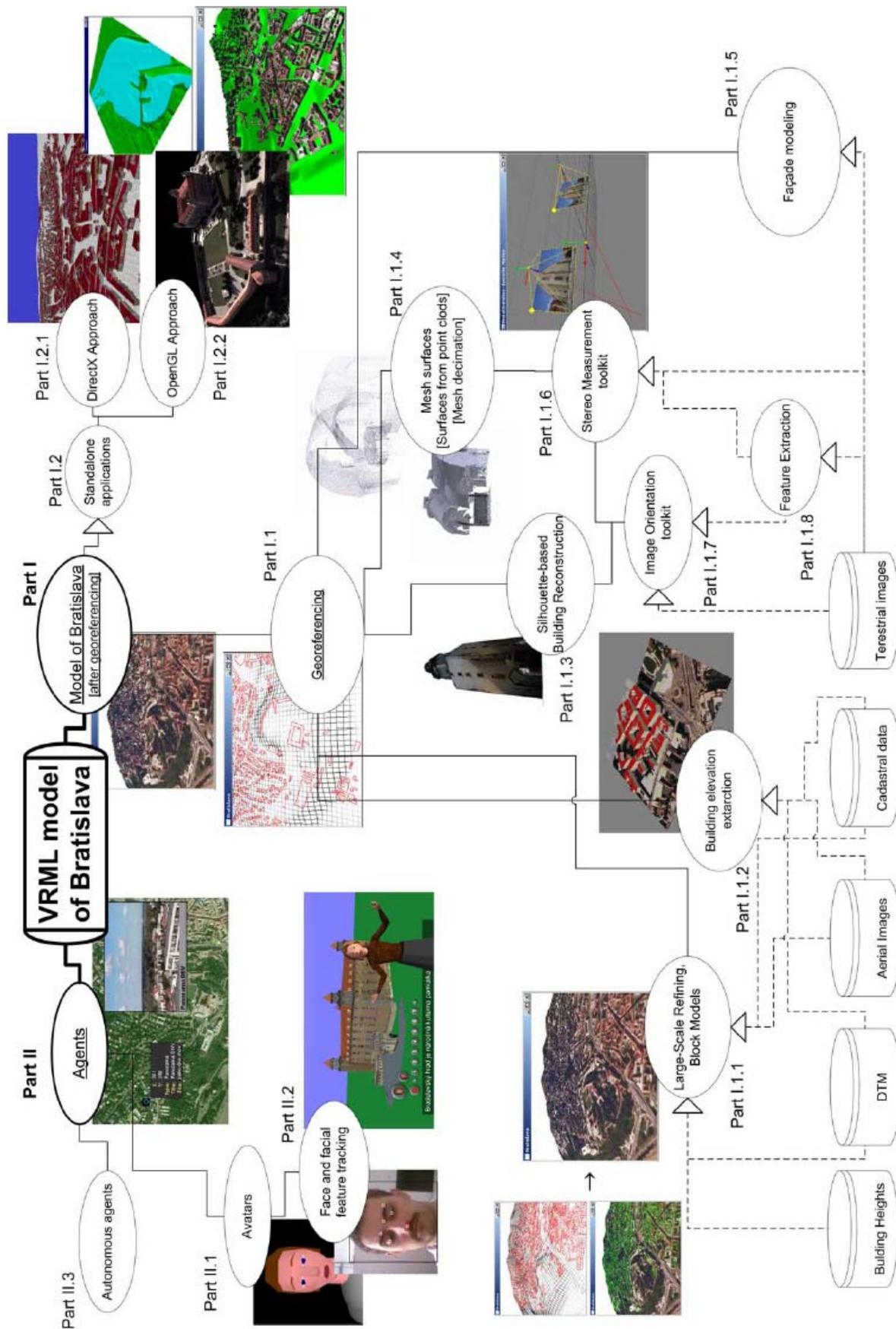


Figure 1. Graph hierarchy of virtual 3D Bratislava project.

Object reconstruction

The main purpose of this work is object reconstruction or simply detection and reconstruction of basic geometric elements as planes (parts of planes), lines or conics. The challenge is – to perform it starting from the pictures of a scene up to reconstruction and mathematical description of 3D elements.

There are lots of articles dedicated to different parts of the solution. This contains image corrections like radial distortion removal, image processing, stereovision – epipolar geometry theory, relevant feature detection and matching, camera calibration and self calibration and so on. A number of minimization techniques is enforced, different approaches have been tested and compared.

The pipeline of such a process can be found in [Pollefeys at all 05]. It presents a complete system that takes a video sequence of a static scene as input and generates 3D model.

Starting from a sequence of images the first step consists of recovering the relative motion between consecutive images. This process goes hand in hand with finding corresponding image features between these images (i.e. image points that originate from the same 3D feature). In the case of video data, the features are tracked until disparities become sufficiently large so that an accurate estimation of the epipolar geometry becomes possible.

The next step consists of recovering the motion and calibration of the camera and the 3D structure of the tracked or matched features. By [Pollefeys at all 05], this process is done in two phases. First, the reconstruction contains a projective skew. This *uncalibrated* approach to 3D reconstruction allows much more flexibility in the acquisition process since the focal length and other intrinsic camera parameters do not have to be measured beforehand and are allowed to change during the acquisition. The obtained reconstruction contains only a sparse set of 3D points. Although interpolation might be a solution, this yields models with poor visual quality. Therefore, the next step consists of an attempt to match all image pixels of an image with pixels in neighboring images, so that these points can also be reconstructed. This task is named dense matching.. Thus, a depth estimate can be obtained (i.e. the distance from the camera to the object surface) for almost every pixel of an image.

By fusing the results of all the

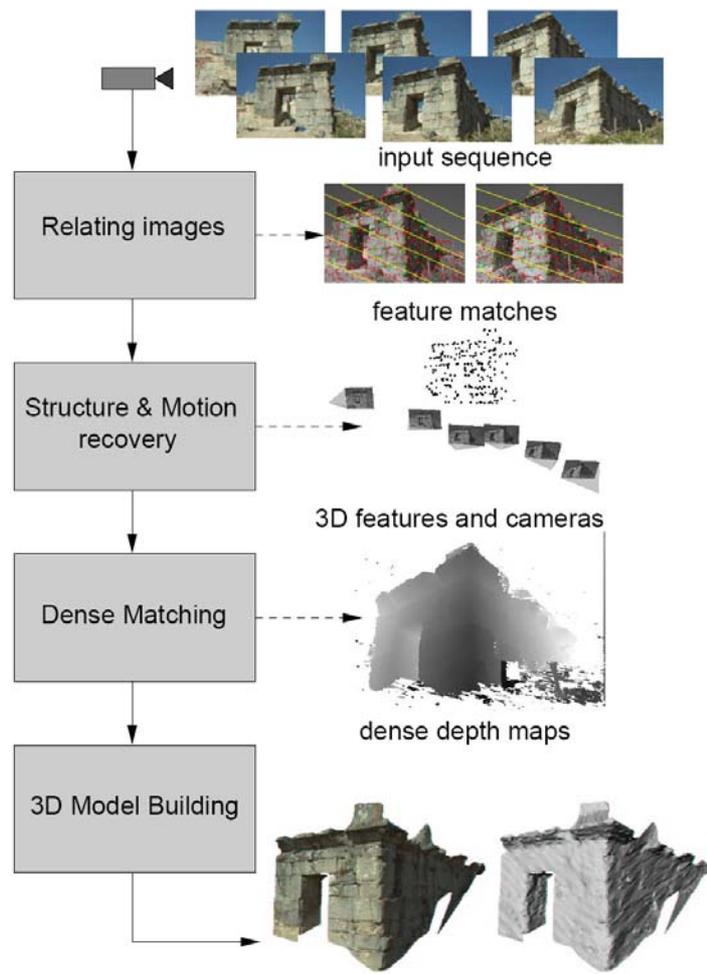


Figure 2. Reconstruction pipeline. [Pollefeys at all 05]

images together a complete dense 3D surface model is obtained. The images used for the reconstruction can also be used for texture mapping so that a final photo-realistic result is achieved. The mentioned steps of the process are illustrated in Figure 2.

The objective of this work differs slightly from the pipeline but individual steps and its succession remains. Several stages are described in following paragraphs in some more detail. First, some background of epipolar geometry is presented. Attention is given to feature detection.

Epipolar geometry

The basis of theory of two-views can be dated to year 1855 when French mathematician Chasles formed the problem of recovering the epipolar geometry from a seven-point correspondence. Eight years later the task was solved by Hesse and in the year 1981 the original eight-point algorithm for the computation of essential matrix was introduced by Longuet-Higgins. The problem of fundamental matrix estimation is studied quite extensively from that time on.

The epipolar geometry is a natural projective geometry between two views of the scene. A view is usually a picture from a camera. The type and the properties of projection are given by construction and adjustment of the camera. The epipolar geometry does not depend on the structure of the scene. It is derived from intrinsic parameters of cameras and their mutual position.

The epipolar geometry arises of two images – projections of the scene. Our goal is to find an equation, which describes the relationship between the pictures. A relation which relates an image point from the first image $\mathbf{x}=[\mathbf{x}_1, \mathbf{x}_2]$ (image coordinates of point, projective coordinates are $[\mathbf{x}_1, \mathbf{x}_2, 1]$) with image point from the second image $\mathbf{x}' = [\mathbf{x}'_1, \mathbf{x}'_2]$ keeping condition that they both are projections of any point \mathbf{X} from the scene is searched. A matrix like this is called Fundamental matrix. It is the algebraic representation of epipolar geometry.

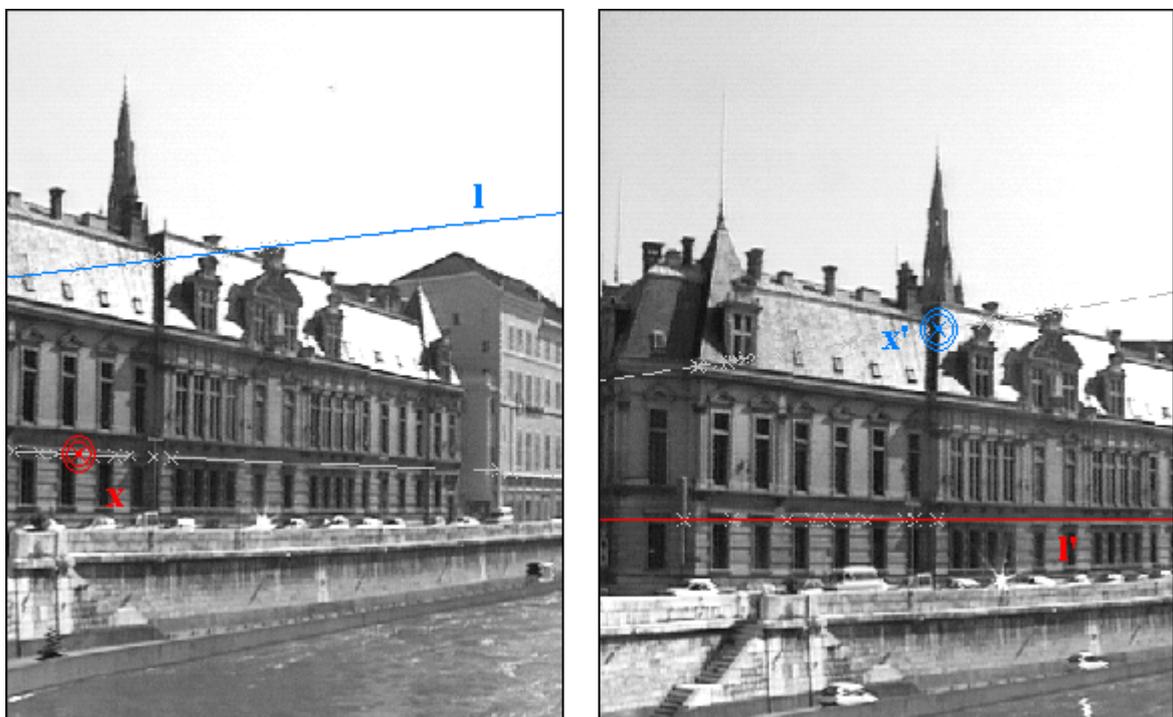


Figure 3. Points and their corresponding epipolar lines. Adopted from [Mohr, Triggs 96].

It can be seen (Figure 3.) that to each point \mathbf{x} in one image a corresponding line \mathbf{l}' in the other image exists – epipolar line. Any point \mathbf{x}' on the line \mathbf{l}' can correspond to \mathbf{x} and simultaneously be an image of point \mathbf{X} in scene. For any pair of valid points \mathbf{x}, \mathbf{x}' exists 3×3 matrix \mathbf{F} – fundamental matrix for which $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$. \mathbf{F} is of rang 2. \mathbf{F} does not depend on scene or choice of selected points, it depends on mutual position of cameras used to take first and second image and on their calibration only.

Fundamental matrix

There exist several methods how to find it. The robust ones work with searching for number of corresponding features in two images and start from this statistically huge set. But this is not possible in our application, now. Our input is a set of several corresponding points from two images and calibration matrix of both cameras (usually the same). Methods used in these situations can be divided to several groups:

- LINEAR algorithm,
- ALGEBRAIC MINIMIZATION algorithm,
- DISTANCE MINIMIZATION

The choice of the applied methods has to consider several requirements. The aim of our implementation is to determine a sufficiently correct method to obtain the fundamental matrix, which is fast for available data. An algorithm for auto-detection of corresponding points was not implemented up to now. The points are assigned by operator (manually).

The 8-point normalized algorithm is a fast and easy to implement method. Usually, it offers quite precise results. It is very suitable as the first step for iterative methods.

If higher precision is required, the algebraic error minimization method is recommended. Similar accuracy is achieved by methods of distance minimization, the highest if the Sampson error is used. It is appropriate as an alternative algorithm.

The 8-point normalized algorithm and algebraic error minimization method were implemented.

Basic linear 8-point algorithm

The best approximation of fundamental matrix is looked for. The equation which defines the fundamental matrix \mathbf{F} is $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$ where \mathbf{x} and \mathbf{x}' is a pair of the matching points in the first and the second image. Their projective coordinates are $\mathbf{x}=(x, y, 1)^T$, $\mathbf{x}'=(x', y', 1)^T$. Each point match results in one linear equation in the unknown entries of \mathbf{F} . For more points matches $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ ($i = 1 \dots n$) linear equations can be stacked up into a matrix. The solution can be found by the least-square algorithm using singular value decomposition. The normalization (simple translation and scaling) of input data is very useful to make the algorithm stable. An important property of fundamental matrix is the singularity. This method, in general, does not produce matrix \mathbf{F} of rank 2.

The algebraic minimization algorithm

The remaining problem is how to guarantee singularity of the constructed fundamental matrix. One possible solution is to construct the singular matrix as a product $\mathbf{F} = \mathbf{M} [\mathbf{e}]_x$ where \mathbf{M} is a non-singular matrix and $[\mathbf{e}]_x$ is any skew-symmetric matrix, with \mathbf{e} corresponding to the epipole in the first image. To guarantee the fundamental matrix properties in such matrix \mathbf{F} , a constraint on \mathbf{F} is added. Matrix \mathbf{F} can be computed from the image point correspondences and known epipole \mathbf{e} by minimization. [Hartley, Zisserman 00] The estimation inaccuracy can be evaluated by an algebraic error $\boldsymbol{\varepsilon}$. It describes a transformation which maps the estimate of the epipole \mathbf{e}_i to the algebraic error $\boldsymbol{\varepsilon}_i : \mathbf{R}^3 \rightarrow \mathbf{R}^8$. The exact epipole is unknown, in reality. We acquire it's estimate using iterative methods. The Levenberg – Marquardt iterative method can be used [Numerical Recipes], [Pollefeys 02]. An estimation of the fundamental matrix \mathbf{F}_0

is calculated using different methods (the 8-point linear algorithm) to get the zeroth approximation of the epipole \mathbf{e}_0 (a right null vector of matrix \mathbf{F}_0). Each iteration aims to change \mathbf{e}_i so that the value $\|\mathbf{e}_i\|$ is minimized.

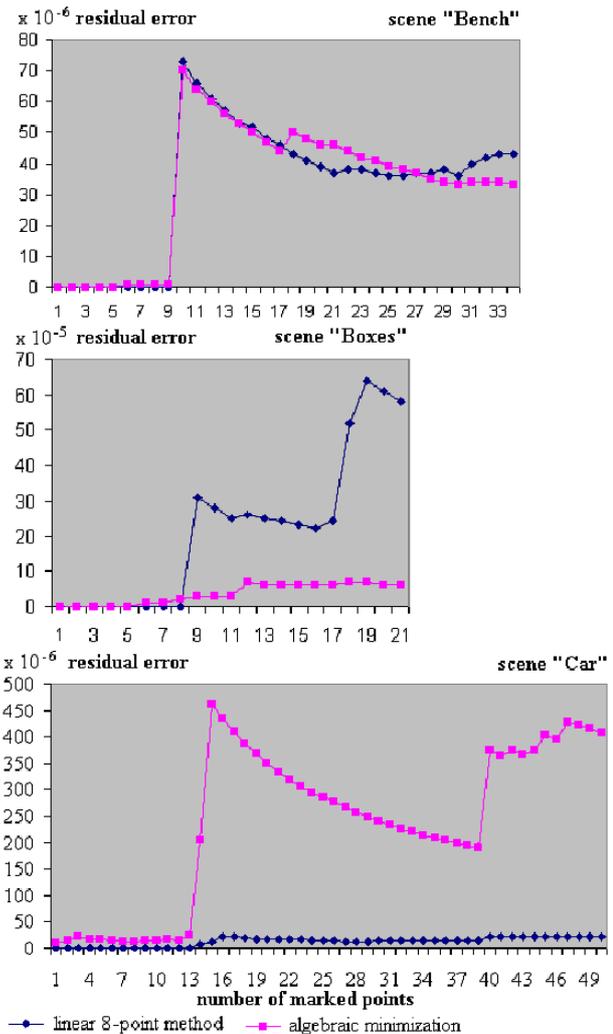
Tests and results

A couple of images with different accuracy and resolution were used to test implemented methods. Synthetic data was used, pictures from tutorial of PhotoModeler [PhotoModeler] and self-made pictures done by a standard camera. Examples of the used scenes are in figure 4.



Figure 4. (top) Test data:
 "Bench" (Res. 280x1024, Foc.l. 6.97mm),
 "Boxes" (Res. 2272x1704, Foc.l. 7.19mm),
 "Car" (Res. 2267x1520, Foc.l. 30.75mm).

Figure 5. (right) Graphs of methods residual error.



To evaluate precision of acquired fundamental matrix the residual error is calculated. The graphs (Figure 5.) show the impossibility to decide positively, which method is more accurate, considering the residual error. In the scenes, which are considered less stable (i.e. where the used matched points are almost coplanar) linear method works better. In scenes defined with higher precision algebraic minimization method is more accurate. Idealized scene was created to understand the methods and their convergence better. Techniques were tested by it (Figure 6.).

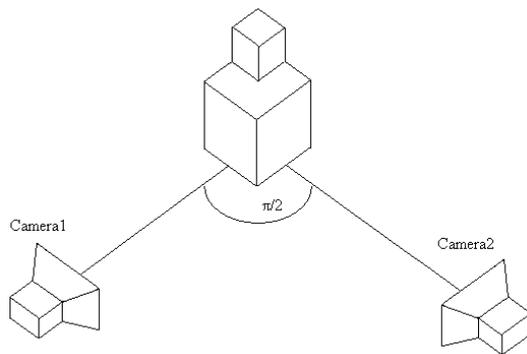


Figure 7. Synthetic scene.

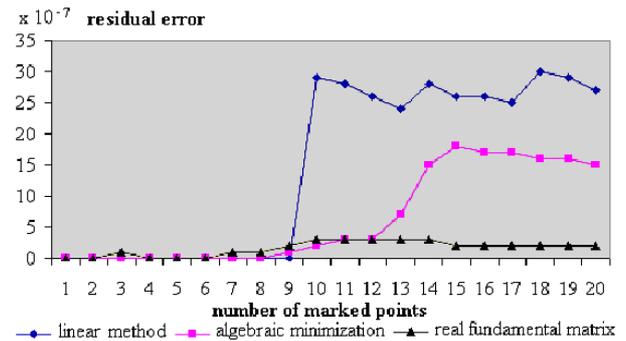


Figure 6. Comparison of residual error of methods for synthetic scene.

The residual error was quantified here for the linear method and algebraic minimization, too. Moreover, exact fundamental matrix derived from the known geometry of the scene was applied (Figure 7). The algebraic minimization seems more suitable by this measurement. All the error values are 10 times lower than in the real scenes. It is caused by much more precise selection of matching points.

The reconstructed points and several predefined faces can be visualized in a 3D scene generated by the OpenGL library. A visualization of the results as a graphical depicting of the reconstructed 3D position of the points in a virtual world is an indirect proof of the calculation accuracy. Such visualization shows how much the reconstruction fits, that is, how the fundamental matrix fits. An experienced operator uncovers in visualization which pair of corresponding points is set incorrectly or improperly.

Automatic computation

When the question of fundamental matrix acquisition using a certain count of corresponding points is solved, another problem arises - how the process could get automated. Two steps should be performed before computation of the fundamental matrix:

- the detection of “points of interest” in both images (it concerns searching for distinctive points which are easy to distinguish in images as corners, edges)
- determination of the possible correspondences (coupling of such eminent points in opposite images, which are similar enough and are projections of single real 3D point with high probability).

Then the conclusions about epipolar geometry are made. The technique of fundamental matrix computation in this case differs from aforementioned methods because of number of available corresponding points.

The robust detection of points of interest constitutes a fundamental step in the characterization and matching of images. The points of interest usually correspond to patterns of significant intensity variation in more than one direction. Many detectors are found in the literature [Laganiere, Vincent 02], [Darilkova 05].

Many candidate feature types have been proposed and explored, including line segments, groupings of edges, and regions, among many other proposals. While these features have worked well for certain object classes, they are often not detected frequently enough or with sufficient stability to form a basis for reliable recognition. The extraction of local features is performed in two steps:

- a) a “where” step involving an interest point detector, and
- b) a “what” step consisting of a local feature extractor [Carneiro, Jepson 03].

The interest point detector selects image locations that contain a high degree of information content, while being robust to common image deformations. The local feature

extractor should provide a representation of such image neighborhoods that is semi-invariant to typical image deformations. Features on the image are automatically deformed with changing a viewpoint. The invariance of descriptor makes them immune against changes in a viewpoint or illumination. Detected features like corners should be invariant under affine transformations like rotation, scaling, intensity changes, independently in each of the images. That offers us a powerful tool of correspondence detection between different views of scene.

Maximally Stable Extremal Regions

Method to detect affine invariant features is a method of Maximally Stable Extremal Regions [Matas et al. 01]. The idea of this method is informally explained in [Matas et al. 01] as follows. Let us assume all possible thresholdings of a gray-level image I through $S=\{0, 1, \dots, 255\}$. We will color all pixels below threshold white and the rest black. The result made of 256 thresholded images, will be a movie. The first image of the movie will be white and then black regions belonging to local intensity minima will appear and these will grow consequently. At some point regions corresponding to two local intensity minima will merge. The last image at the end of the movie will be black. The set of all connected components of all frames of the movie is the set of all extremal regions. Afterwards, we will select these extremal regions which support is virtually unchanged over a range of threshold. The selected regions were designated as Maximally Stable Extremal Regions (Figure 8). The first step to find MSERs is sort pixels by intensity. Then algorithm goes through the image intensities from 0 to 255 and creates tree structure (starts by leaves) using several properties. The tree describes the image. MSER detection step is a bottom-up walk through the tree and field of leaves.

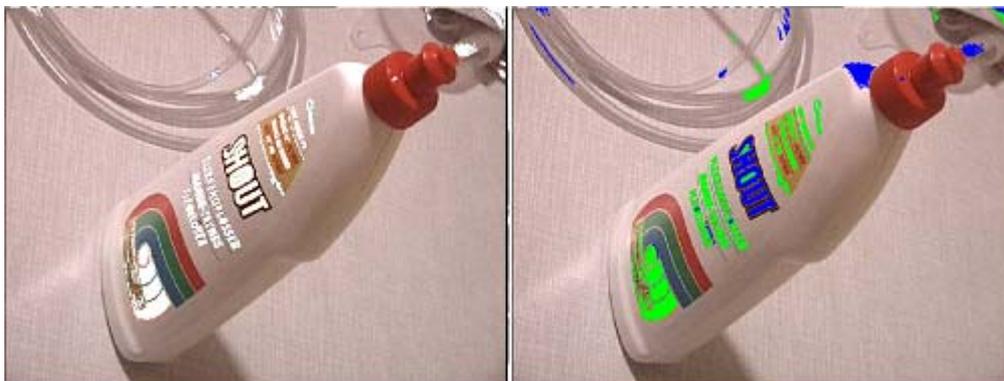


Figure 8. The detected MSERs. Comparison of Dr. Matas' team results (left) and our implementation output (right). The results are identical by 95%.

Local affine frames

The next step after MSER region detection is to find Local Affine Frames (Local Affine Frames - LAF). It can be viewed as local coordinate system, which is detected invariantly to both affine transformations (geometric and illumination). Local affine frames are used to provide normalisation of image patches into canonical frame to enable direct comparison with Intensity Normalised Cross-Correlation Method. LAF is the set of three points, which define the local coordinate system. These three points need to be affine invariant. The first type of LAF is obtained from covariance matrix. From this matrix, we obtain properties of ellipse E which approximates the detected MSER (a center of the ellipse and two ellipse axes define LAF) [Jancosek 05]. E can be by transformation normalized to unit circle. It transforms the local coordinate system defined with ellipse parameters to new global, but this normalizes

the MSER region up to a known rotation. Thus, we have to complete the affine frame to resolve the rotation ambiguity. Seven different directions were used in this work (Figure 9).

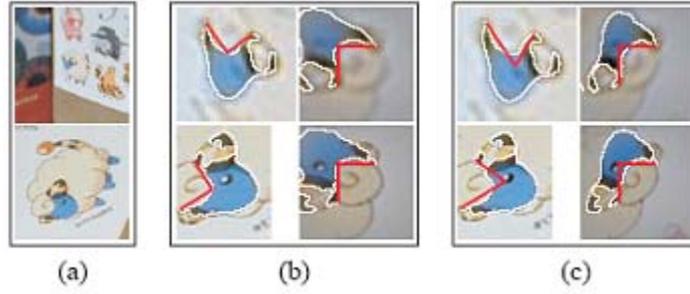


Figure 9. a) Two original images, b)c) MSER region detected in the image (left) using different orientation definition and normalised (right).

Correspondence

Tentative correspondence estimation is based on intensity-normalized cross-correlation [Barnea, Silverman 72] between two normalized regions in terms of the local coordinate frames.

$$NCC(I_1, I_2) = \frac{\sum_x (I_1(x) - \bar{I}_1)(I_2(x) - \bar{I}_2)}{\sqrt{\sum_x (I_1(x) - \bar{I}_1)^2 \sum_x (I_2(x) - \bar{I}_2)^2}},$$

where $\bar{I}_1 = \frac{1}{N} \sum_x I_1(x)$, $\bar{I}_2 = \frac{1}{N} \sum_x I_2(x)$ are the means of windows I_1 and I_2 . The cross correlation takes on values in $[-1, 1]$ (1 being the most similar, -1 being the least similar) and is invariant to illumination transformations such as contrast and brightness modifications.

In general, if we put into NCC two MSERs without their surroundings, the MSER has a pure histogram and the result will be delusive. So we need to get MSER with some surroundings.

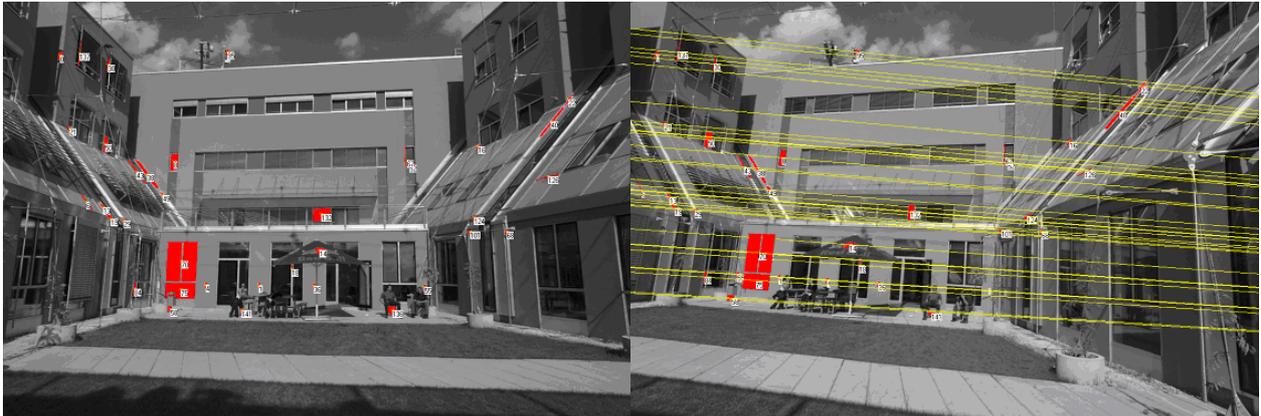


Figure 10. A red numbered areas are corresponding MSERs confirmed by fundamental matrix test. Yellow lines are appropriate epipoles.

True tentative correspondences

In this work a new method called True Tentative Correspondences (TTC) is described. This method is based on two observations. The first one is a powerful rule, The Sideness Constraint [Ferrari, Tuytelaars 03], which is not most commonly used in other methods (Figure 11).

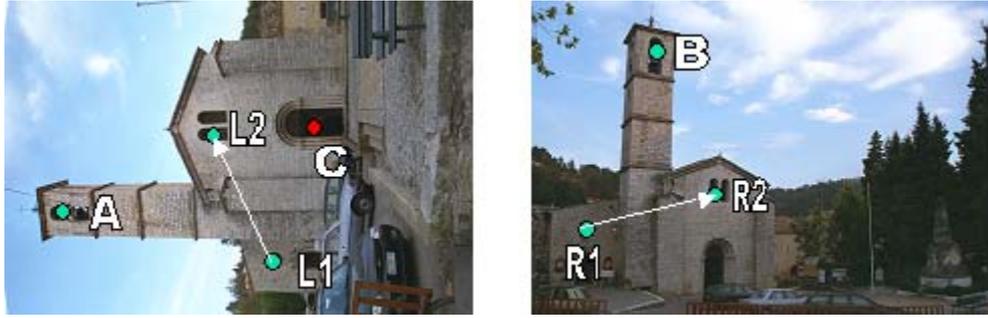


Figure 11. The Sideness Constraint.

The idea is as follows: If the corresponding directed lines are known, then we can divide the left image into the left and the right part. If there is a correspondence pair of points A, B, they have to lie on the same part of the left and the right image (see Figure 11). The second one is our idea of solving rotational ambiguity of Local Affine Frames (mentioned before). As the rotation we have chosen the direction to the center of gravity of the former detected MSER. This MSER is the MSER from the detected tentative correspondence pair we assume to be true one. The essential idea of our method is to search for the following correspondence using the best correspondences found before.

Step 1: We take the correspondence regions A_0, B_0 with the best correspondence value computed in a comparative process on LAF types with rotation obtained from MSER boundary curvature analysis. In every experiment we have realized that the best correspondence is the geometric one. Then for each region in the left image we compute LAF with direction to A_0 and for each region on the right image with direction to B_0 .

Step 2: The comparative process is running again, then we get the first best different correspondence regions A_1, B_1 with the best correspondence value different from A_0, B_0 which are not closer to A_0, B_0 than 10 pixels on both images separately.

Step 3: There are two tentative correspondences and we assume that they are geometric and we can now use The Sideness Constraint. Then for each region in the left image we compute two LAFs with directions to A_0, A_1 and for each region on the right image with directions to B_0, B_1 . Afterwards we compute a correspondence map and save it to memory.

Comment: We reject correspondences by The Sideness Constraint by setting the correspondence value to a certain big number (5.0) in the correspondence map of regions which are on the different sides of the corresponding directed lines.

Step 4: This step is repeated until $n=7$:

1. Load correspondence map and reject correspondences by The Sideness Constraint for each pair of correspondences from the True Correspondences Set $\{A_0, B_0; A_1, B_1; \dots; A_n, B_n\}$ by changing the correspondence map

2. Run comparative process on this correspondence map and add the first best different correspondence to the True Correspondences Set

To get the set of TTC we have developed an algorithm, which saves TTC in a binary tree of height 8, where every path from the root to the leaf is one TTC. There are 8 correspondences as output from this method, which we named True Tentative Correspondences (TTC).

Used type of rotation is invariant under change of viewpoint (in perspective image pair) only if, the point to which LAF is rotated, lies on the same plane as the detected MSER (local planarity assumption) to which LAF belongs. In many cases this error is smaller than error

arose by MSER local affine frame estimation. Our experiments show superior performance of LAF with this type of rotation (Figure 12).



Figure 12. Eight correspondence regions from TCC method (blue).

Conclusion

A number of approaches exist in different stages of multiple view object reconstruction process. Two of them are inspected in this work but the list is not complete. The differences can be found between pipeline of stereo reconstruction and the one presented in part Object reconstruction (which performs the reconstruction from video sequences) but methods are similar and the goal is the same – 3D model of the scene.

Two images of the scene are input to the method. In former part was shown, how maximally stable regions can be found in images. The new true tentative correspondence algorithm for finding tentative correspondences between MSER regions is presented and there is also experimentally shown, that in every case there exist at least 8 correspondences in TTC tree. We need only 8 geometric correspondences to compute epipolar geometry between two images.

The part Epipolar geometry is dedicated to comparison of two methods for acquisition of the fundamental matrix by 8 points marked in two pictures of a scene. Linear method and the method of algebraic minimization were exploited. The presented comparisons show similarity of the methods results. Well-defined scenes have significantly better results with the fundamental matrix improved by algebraic minimization starting from a matrix given by the linear method. If the scene is described by improper set of points, the linear method is more suitable. It is possible to recognize this case by monitoring of several properties of computation and of partial results. The choice of another set of marked points can be more efficient step to get the accurate fundamental matrix.

Standard camera suffices to take two pictures of any scene and to reconstruct its geometry. Some of camera parameters are required to be known (published usually by the producer). Preliminary tests confirm possibility of methods combination. More tests, algorithms adaptation and time efficiency improvement is the task of future work. The tests tasks are: Is the eight point groups computed by TTC congruous for precise fundamental matrix computation? Is automatic selection of best group possible?

The combination of presented methods shows a new way of 3D scene reconstruction. Two images of a scene are input to our method. Standard camera suffices to take the pictures. Some of camera parameters are required to be known (published usually by the producer). The output is reconstruction of scene and cameras geometry and 3D position of any point detected as correspondence in the input images. The combination of presented methods and results of dense aggregation of image points transforms input images to point cloud of

reconstructed 3D points of surfaces. This can be used as input to another group of algorithms like fast rendering techniques or mesh simplification.

This method combination provides a way, how the slow stochastic RANSAC method can be substituted with use of the outputs from the deterministic TTC method and fast fundamental matrix computation and precision validation.

Acknowledgements

This work is the part of project supported by VEGA grant Virtual Environments for WWW No. 1/0174/03. Our thanks to Doc. RNDr. Andrej Ferko, PhD. (FMFI – UK, Bratislava), Dr. Ing. Jiří Matas, Ing. Michal Perďoch, Ing. Štěpán Obdržálek (CMP - The Center for Machine Perception, Praha), Joachim Bauer Dipl.-Ing. (VRVis, Graz) for willing help and experimental data they had gave to us.

References

- [Barnea, Silverman 72] D. I. Barnea, H. F. Silverman, *A class of algorithms for fast digital image registration*. IEEE Trans. Computers, 21, pp. 179-186, 1972
- [Carneiro, Jepson 03] Carneiro, G.; Jepson, A.D.: *Multi-scale Phase-based Local Features*, proceedings of the Conference on Computer Vision and Pattern Recognition, Volume I, pp. 736, June 2003.
- [Darilkova 05] Darilkova, K.: *Reconstruction and stereovision*, Project of dissertation, FMFI Comenius university, Bratislava, Slovakia, 2005.
- [Ferrari, Tuytelaars 03] V. Ferrari, T. Tuytelaars, L. Van Gool, Wide-baseline multiple-view correspondences. *in Proc. IEEE. CVPR*, 718-725, 2003.
- [Hartley, Zisserman 00] Hartley, R.; Zisseman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, United Kingdom, ISBN 0 521 62304 9, 2000.
- [Jancosek 05] Jančošek, M.: *Feature Detection and Tentative Correspondence Estimation in Wide Baseline Stereo*, Master thesis, FMFI Comenius university, Bratislava, Slovakia, 2005.
- [Laganieri, Vincent 02] Laganieri, R.; Vincent, E.: *Wedge-Based Corner Model for Widely Separated Views Matching*, proceedings of the 16 th International Conference on Pattern Recognition, Volume 3, pp. 30672, August 2002.
- [Matas et al. 01] J. Matas, O. Chum, M. Urban, and T. Pajdla. *Distinguished regions for wide-baseline stereo*. Prague: Center for Machine Perception, Research Report CTU–CMP–2001–33, K333 FEE Czech Technical University, November 2001
- [Mohr, Triggs 96] Mohr, R.; Triggs, B. 1996. *Projective Geometry for Image Analysis*. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MOHR_TRIGGS/node50.html, (5.10.2004)
- [Numerical Recipes] Numerical recipes in C: the art of scientific computing, Cambridge University Press, ISBN 0-521-43108-5
- [PhotoMod] PhotoModeler Pro 5.0 Demo, Copyright 1993 – 2002 Eos Systems Inc. <http://www.photomodeler.com/ddlp.html>, (8.9.04)
- [Pollefeys 02] Pollefeys, M.: *Visual 3D Modeling from Images*, Tutorial Notes, 2002, University of North Carolina - Chapel Hill, USA, www.cs.unc.edu/~marc/tutorial/node160.html
- [Pollefeys et al. 05] Pollefeys, M.; Van Gool, L.; Vergauwen, M.; Verbiest, F.; Cornelis, K.; Tops, J.: *Video to 3D*, Photogrammetric Computer Vision, ISPRS Commission III, Symposium, September 9 - 13, 2002, Graz, Austria, part 3A/B, pp. A-252, ISSN 1682-1750